Chapter 3

An extensive screen of the HLA region reveals an independent association of HLA class I and class II with susceptibility for systemic lupus erythematosus

H.A. Martens[1*], I.M. Nolte[2*], G. van der Steege[3], M. Schipper[3], C.G.M. Kallenberg[1], G.J. te Meerman[3], M. Bijl[1]

[1] Department of Rheumatology and Clinical Immunology, [2] Department of Epidemiology, [3] Department of Medical Genetics, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands

[*] Authors contributed equally

Published in the *Scandinavian Journal of Rheumatology 2009; 38: 256-262*

Abstract

Objective. The association of systemic lupus erythematosus (SLE) with the human leukocyte antigen (HLA) region is well known. In this study, we analyzed the involvement of the HLA region in susceptibility for SLE in a stable founder, Caucasian population of SLE patients.

Methods. We performed an extensive screen of the entire HLA region on 103 SLE patients and family-based controls. Both single locus and haplotype sharing analysis were performed. The additional disease locus test (ADLT) was applied to examine the nature of the observed associations and to distinguish true causal associations from associations due to linkage disequilibrium (LD).

Results. Significant associations were observed at markers in the HLA class I, class II and class III regions by using haplotype sharing and single locus methods. The haplotype sharing methods revealed the most significant difference at marker D6S1666 in the HLA class II region (p_{HSS}=0.00037; p_{CROSS}=1.7 $\cdot 10^{-5}$). The most significant result of single locus association was shown at marker D6S265 in the HLA class I region (p=0.00019). The ADLT demonstrated that these markers represent independent associations.

Conclusion. HLA class I, class II and class III are associated with SLE, but only class I and class II contribute independently to increased risk of SLE.

Introduction

Systemic lupus erythematosus (SLE) is a complex autoimmune disease with a broad spectrum of clinical manifestations characterised by the production of a multitude of autoantibodies against self-antigens. The exact pathogenesis of SLE is unknown, but genetic predisposition is thought to play an important role.[1;2] Several genes have been reported to be associated with SLE, including genes localized on the human leukocyte antigen (HLA) region on the short arm on chromosome 6.[1-9] Certain HLA alleles, for instance the HLA-DR2 and DR3 alleles, are associated with a two to threefold risk for the development of SLE.[10] A genome scan meta-analysis for SLE susceptibility loci revealed the most significant result on chromosome 6p22.3-21.1, containing the HLA region. This result was found for all genome scans independent of race, but was more prominent in the analysis of Caucasian populations alone.[11]

The exact role of the HLA molecules in the pathogenesis of SLE is not clear, but they do have important functions in the immune system. Several autoimmune disorders (e.g. type 1 diabetes, rheumatoid arthritis) are associated with HLA alleles. The HLA class I region, containing the HLA A, B, and C genes, codes for HLA molecules that play a role in presenting antigenic peptides to CD8+ cytotoxic T cells. The class II region contains the DR, DQ and DP genes which are expressed on antigen presenting cells. The association between HLA class II genes and SLE proved to be more consistent then the association between class I genes and SLE. HLA class III genes encode for amongst others complement C4, C2 and tumor necrosis factor (TNF)α.

In this study, we analyzed the involvement of the HLA region on chromosome 6 in susceptibility for SLE in a stable founder population of SLE patients and their relatives from the northern part of the Netherlands. We genotyped 19 microsatellite markers spanning the entire HLA region. The use of multiple densely placed microsatellite markers allowed us to perform not only single locus association analysis, but also haplotype analysis. To examine the nature of the observed associations, the additional disease locus test (ADLT) was applied to discriminate true causal associations from associations due to LD.

Methods

Subjects

In the period from November 2000 till November 2001, all SLE patients from the SLE cohort at the University Medical Center Groningen were asked to participate in a study on genetic predisposition of SLE. All patients met the criteria for SLE according to the American College of Rheumatology (ACR).[12] Out of 164 SLE patients approached, 107 (65%) agreed to take part in the study. Four patients who were not of Caucasian descent were excluded to avoid influence of ethnicity, thereby yielding 103 patients who could be studied for genetic susceptibility for SLE (table 1).

Table 1: basic characteristics of 103 SLE patients

Median age, years (range)	43 (23-78)
Sex female, n (%)	89 (86%)
Median age of onset, years (range)	31 (8-73)
Median duration of disease, months (range)	131 (21-516)
Race, n (%)	
Caucasian	103 (100%)
ACR criteria, n (%)	
Malar rash	37 (36)
Discoid rash	31 (30)
Photosensitivity	52 (50)
Oral ulcers	13 (13)
Arthritis	67 (65)
Serositis	39 (38)
Renal disorder	42 (41)
Neurologic disorder	7 (7)
Hematologic disorder	75 (73)
Immunologic disorder	91 (88)
anti-dsDNA antibodies	81 (79)
anti-Sm antibodies	13 (13)
anti-phospholipid antibodies	21 (20)
Anti-nuclear antibody	103 (100)

Family members, ideally both parents or spouse and a child, were invited by the participating patients to join in this study. All family members were unaffected. Either the non-transmitted chromosomes of the parents or the chromosomes of the spouse served as controls. Other family members could be used for linkage phase determination, in the case that neither both parents nor spouse and child were participating. The study was approved by the local medical ethics committee.

Genotyping

DNA was either extracted from 20 ml ethylenediaminetetraacetic (EDTA)-blood following standard procedures or from serum by using the QIAamp DNA Midi Kit (Qiagen, Valencia, Cam, USA) and was stored at -80°C. Subsequently, 19 polymorphic microsatellite markers were genotyped in all patients and family members. Table 2 presents the marker and primer and probe information in detail.

The reaction volume was 10 µl, which included approximately 25 ng DNA. For each polymerase chain reaction (PCR), 0.5 units *Taq* DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden) was used to amplify the fragments. Reaction mixtures contained 0.2 mM dNTP (Roche Diagnostics, Mannheim, Germany), 2.5 mM $MgCl_2$, 10 mM Tris-HCl, 50 mM KCl (Amersham Pharmacia Biotech) and 0.25 µM of each primer (with one primer 5' labeled with a fluorochrome 6-FAM, HEX (Sigma, Malden, the Netherlands) or NED (Applied Biosystems, Foster City, USA). Cycling was performed on a PTC-225 thermal cycler (MJ Research, Waltham, USA) and a PrimusHT (MWG Biotech, Ebersberg, Germany). A standard protocol was used for amplification. Post-PCR multiplexing was performed by combining 2-10 µl (based on signal strength) of the PCR products. A 2.3 µl sample of the pooled fragments was mixed with 2.5 µl MilliQ and 0.2 µl ET-400R size standard (Amersham Pharmacia Biotech) and separated on a MegaBACE 1000 capillary sequencer (Amersham Pharmacia Biotech) according to the manufacturer's protocol. Results were analyzed using genetic profiler version 2.0 (Amersham Pharmacia Biotech). The assignment of the genotypes was blinded for affection status (patient or control) and family structure.

Statistical analysis

After genotyping the 19 microsatellite markers in all participants, they were first tested for quality by the Hardy-Weinberg equilibrium test. Among healthy individuals, the observed genotype frequencies were compared with

frequencies expected from the allele frequencies by a chi-square test with the number of degrees of freedom equal to the number of alleles minus 1. Markers that failed this test (i.e. p-value<0.05/19; significance after Bonferroni correction), were excluded from the data set for further analysis. As a test of LD to prove the presence of conserved haplotypes, the D' for multi-allelic markers was used. The significance of the observed D' value was assessed by randomization. For each locus, the observed alleles were randomly distributed over the haplotypes. The significance of the observed D' value was determined by the fraction of 10 000 randomizations that revealed a larger D'.

Table 2: characteristics of the 19 microsatellite markers used in this study for analysis of the HLA region. Primer sequences are as in Diepstra et al.[28]

Marker	Heterozygosity	HWE[a] p-value
6SL001	0.905	0.93
6SL002	0.647	0.99
DNRNGCA	0.808	0.93
D6S2658	0.590	0.12
6BO001	0.684	0.96
D6S2444	0.682	0.98
G511525	0.776	0.48
D6S1666	0.860	0.22
D6S2665	0.831	0.89
D6S2670	0.890	0.94
D6S273	0.792	0.99
D6S2671	0.785	0.99
MICA	0.596	0.82
D6S2673	0.789	0.23
D6S2694	0.912	0.53
D6S2700	0.718	0.88
D6S265	0.683	0.51
D6S2707	0.854	0.50
D6S1683	0.706	0.71

a Hardy-Weinberg equilibrium; p<0.05/19 defines significant deviation from HWE and is regarded as evidence for low genotype quality.

For the trios in our data set, a set of patient and a set of control haplotypes were determined. When DNA from parents was available for haplotype phase determination of the alleles, the non-transmitted haplotypes from the parents were used as controls. If DNA was obtained from a child and a spouse, both haplotypes of the spouse were regarded as controls. When only one family member (a parent, child or sib) was available, the haplotype not present in the patient was used as a control haplotype. Hence the number of alleles and haplotypes included in the analysis is less than twice the number of genotypes. If no family members were willing to participate, the patient was still included but his/her genotypes were not used for haplotype analysis, since analyzing them would introduce bias due to unbalanced phase information between cases and controls.

The two haplotype sets were compared using the haplotype sharing statistic (HSS) and the CROSS test, two methods that analyze the length of haplotype similarity between patients and controls. The validity of these methods has been demonstrated previously both in simulation studies as well as in empirical data including one study on the same population as used here, analyzing chromosome 1 candidate region for SLE susceptibility.[13-16] In short, the HSS and the CROSS tests both assume that haplotype segments of patients are conserved in the region spanning a disease locus. By contrast, control individuals are not expected to have conserved haplotypes centered at a particular locus. Therefore, patients are considered to display excess of haplotype sharing, which is defined as the number of consecutive overlapping intervals between haplotypes, as compared to controls. The HSS compares the mean length of haplotype sharing among patient haplotypes with that among control haplotypes. The CROSS test hypothesizes that patient and control haplotypes are different from eachother at markers close to the disease locus and hence that haplotype sharing of a patient and a control haplotype is less than haplotype sharing of two patient or two control haplotypes.

In addition to the HSS and the CROSS, classical association analysis was performed by comparing the allele frequencies of each marker between patients and controls using a chi-square test for multiple alleles. Odds ratios (ORs) were calculated for the allele of interest versus all other alleles and 95% confidence intervals (CIs) are determined using a Bonferroni multiple testing correction for the number of alleles. The multi-allelic transmission disequilibrium test (TDT) was applied to the trios consisting of a patient and both parents.[17;18]

To determine which associations were causal and which were caused by LD

with a causal locus, we applied the ADLT.[19] The ADLT is a permutation test on the patient and control haplotypes conditioned on the allele present at the supposed disease locus. In this way the permutation procedure leaves the association at the disease locus unharmed and measures only additional causal association at the other loci. If a locus still shows a significant difference between patients and controls after the permutation, it means that this association cannot be explained by LD with the supposed disease locus and therefore is an independent source for disease susceptibility.

Results

From the 103 trios, we derived 206 patient haplotypes and 171 control ones. Of these 171 control haplotypes, 17 came from an incomplete trio and hence, 154 could be used to construct 77 (pseudo-)control genotypes. Fifty-six trios were classical trios of a patient and both parents and were used for the TDT. Strong LD was observed over the entire region, but one complete LD block seemed to be present from marker D6S2444 to D6S1683 (figure 1).

Fig. 1

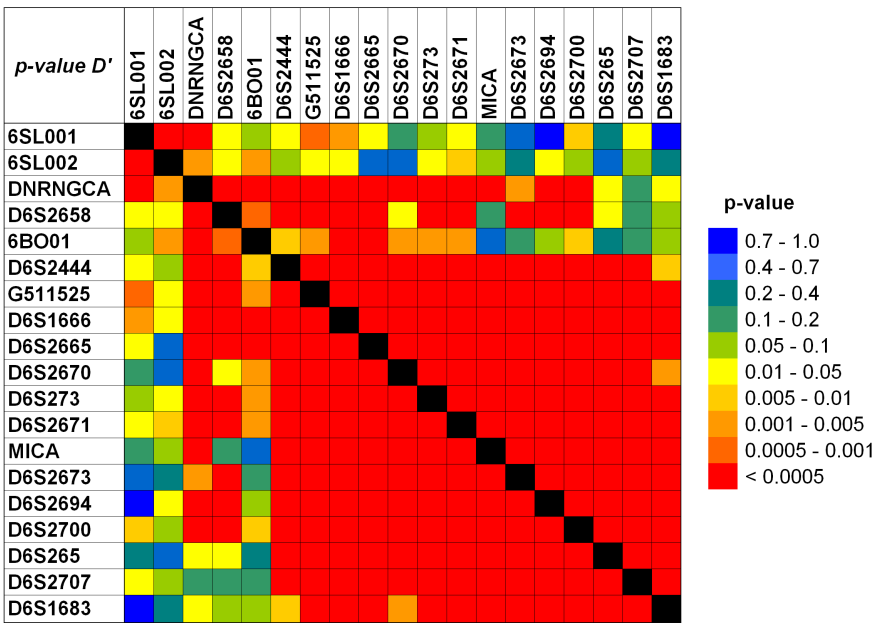

Fig. 1: Linkage disequilibrium as represented by the significance of D'. For each pair of markers, the multi-allelic D' was calculated and its significance was determined by means of a permutation procedure (see Methods).

Table 3: allele frequencies for the most significant markers in single locus association

Marker	Allele	freq. patients	freq. control	p-value[a]	OR[b]	95% CI[c]
D6S2658	352	26.3%	38.6%	0.024	0.29	0.32-1.01
D6S1666	158	1.1%	9.7%	0.0025	0.11	0.01-0.89
MICA	184	66.9%	50.7%	0.013	1.97	1.09-3.56
D6S265	132	6.7%	19.6%	0.00019	0.29	0.11-0.74

a Global p-value calculated by chi-square test on all alleles with expected count > 5 in both the patient and the control group. **b** OR = Odds ratio for this allele versus all other alleles **c** 95% CI = 95% Confidence interval corrected for the number of alleles with expected count > 5 in both the patient and the control group.

To investigate the involvement of the HLA region in our SLE population, haplotype analysis was performed using the HSS and CROSS test. Both the HSS and the CROSS methods revealed significant differences between patients and controls for many markers over the entire HLA region (figure 2).

Fig. 2

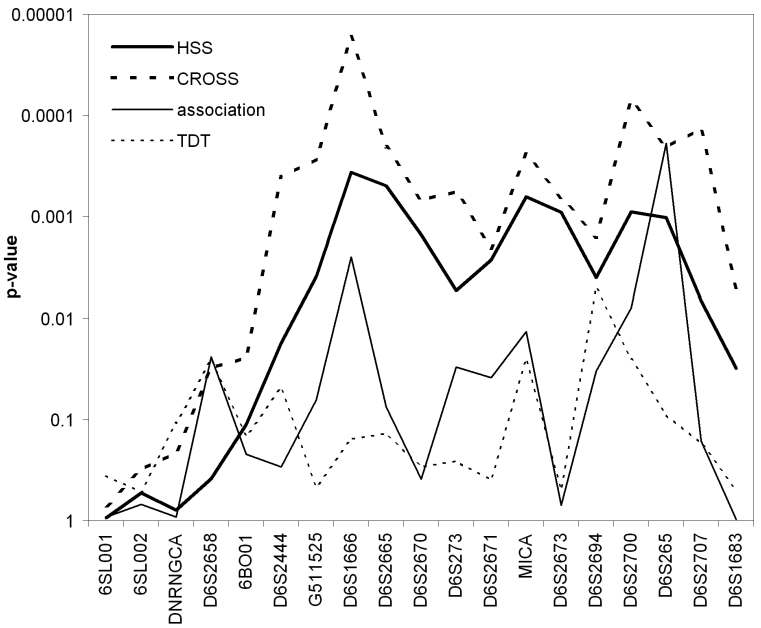

Fig. 2: P-values of the HSS (thick solid line), the CROSS test (thick dotted line), single locus allelic association analysis (thin solid line) and the TDT (thin dotted line). Markers in the HLA class I, class II and class III regions yield significant results.

The most significant p-values were at markers D6S1666 in HLA class II (p_{HSS}=0.0016; p_{CROSS}=1.7 ·10^{-5}), MICA in HLA class I (p_{HSS}=0.0015; p_{CROSS}=0.00024) and D6S2700 in HLA class I (p_{HSS}=0.0051; p_{CROSS}=7.1 ·10^{-5}).

Fig. 3

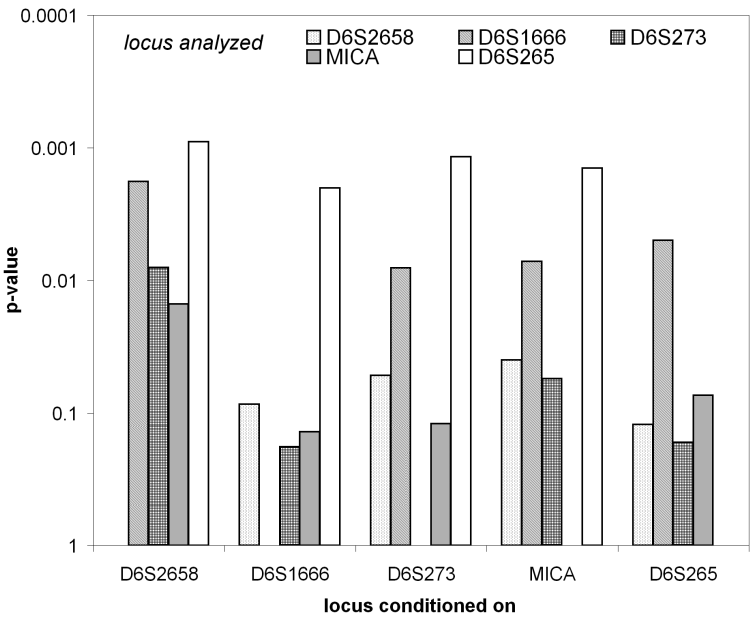

Fig. 3: Result of ADLT that tests for independence of association results. The markers D6S1666 in the HLA class II region and D6S265 in the HLA class I region represent independent associations as their associations remain significant after conditioning. The associations at D6S2658 and MICA appear after conditioning on D6S1666 or D6S265 implying that these are likely caused by linkage disequilibrium with these loci.

Single locus allele association analysis confirmed the results of the HSS and CROSS test (D6S1666: p=0.0025; MICA: p=0.013; D6S265: p=0.00019, table 3). The frequency of the most significant allele 158 at D6S1666 was 9.7% among controls and 1.1% among patients (OR 0.11; 95% CI 0.01-1.01). Allele 184 of the MICA marker was more often present in patients than in controls (66.9% vs. 50.7%; OR 1.97; 95% CI 1.08-3.60). At D6S265, allele 132 was present in 19.6% in the controls and only 6.7% in the patients (OR 0.29; 95% CI 0.11-0.80). An additional association was observed at marker D6S2658 (global p=0.024; allele 352 OR 0.57; 95% CI 0.31-1.03).
Finally, TDT was applied on trios consisting of the patient and both parents. Despite the low number of trios, the TDT was most significant at marker

D6S2658 (global p=0.026; transmission frequency (TRF) of allele 350=64.7%, TRF of allele 352=32.6%) in class II, MICA (global p=0.025; TRF of allele 184 =68.1%) in class I and D6S2694 (global p=0.0051; TRF of allele 277=68.0%; TRF of allele 309=14.3%) in class I.
To discriminate between causal association and association due to linkage disequilibrium, ADLT was performed (figure 3). The ADLT revealed that the association at D6S1666 is not enough to explain the association at D6S265, and vice versa, as after conditioning on either loci the association at the other did not disappear. However, the associations observed at D6S2658 and MICA were not significant after conditioning on D6S1666 or D6S265 implying that these associations are likely caused by LD with one or both of the D6S1666 or D6S265 markers.

Discussion

This study showed that the entire HLA region is strongly involved in the susceptibility for SLE in a stable founder, Caucasian population. Both single locus association and haplotype sharing analysis revealed three regions that were significantly associated. Associations were also found for all HLA classes using TDT, but these were less significant than those found using HSS and CROSS. This was caused by the lower number of trios analyzed using the former test. However, the results of the TDT were still significant, so firm conclusions can be drawn. After additional analysis, only HLA class I and II seemed to contribute independently to the risk of SLE and the association of HLA class III was probably caused through LD with HLA classes I and II.
The strong association of SLE with the HLA region is already well known.[2-8] In a genome scan meta-analysis performed in 2005, the most significant association with SLE was found for the HLA region, especially in Caucasian subjects.[11] In addition, in other autoimmune disorders, for example type 1 diabetes and rheumatoid arthritis, an association with HLA alleles has been demonstrated, indicating a major role for HLA molecules in susceptibility for autoimmune disease.
Because we used the ADLT to examine the true nature of multiple associations in a region, we were able to determine that only HLA class I and II were independently associated with SLE. A strong association between HLA class II alleles and SLE was demonstrated by McHugh et al. in 2006.[20] One of the suggested pathogenetic mechanisms of HLA class II molecules in SLE is their contribution to activation of CD4+ T cells. These cells are required for somatic hypermutation, and extensive antigen-driven somatic hypermutation was

observed in B cells of humans and mice with lupus.[21] Furthermore, it has been suggested that HLA class II molecules may play a role in presenting peptides derived from apoptotic cells.[4] As apoptosis is thought to be an important pathogenic mechanism in SLE, HLA class II molecules may contribute to the pathogenesis of SLE. In addition, certain combinations of HLA DR alleles were shown to be associated with increased antibody production in SLE patients as well as in unaffected family members.[22] The association between class I alleles and SLE, however, is not a consistent finding in literature. Only a few studies showed an independent association of HLA class I with SLE.[5;23]

In addition, the HLA class III region has been reported to be associated with SLE. This region encodes for complement proteins and TNFα. Deficiency of the proteins coded by the C4A, C4B and C2 genes lead to SLE.[24] Several studies reported an association of C4 with SLE (reviewed in[25]). Individuals who are homozygous for the C4AQ*0 null allele are at an increased risk of 9.7-16.9 of developing SLE, while those heterozygous still have an increased risk of 2.3-4.9.[25] One study reported an association of heterozygosity for the C2Q*0 null allele with SLE, but others could not confirm this. TNFα levels are increased in SLE patients and strongly correlate with SLE activity.[26] In addition, a meta-analysis on 19 association studies demonstrated association of the promoter polymorphism -308A/G in TNFα only in the European population with a twofold risk of SLE when carrying the A allele and threefold when carrying it homozygously.[27] In our study, an association between SLE and HLA class III was also found. However, ADLT showed that this association is probably due to LD.

As this study was designed to elucidate the nature of the various known SLE risk loci in the HLA region, we do not have any information about the specific HLA alleles. However, the extensive screen of the entire HLA region and the use of ADLT gave us the opportunity to better pinpoint the true causal associations. In addition, the inclusion of family-based controls and the use of multiple densely placed microsatellite markers allowed us to perform not only single locus association analysis but also haplotype analysis (HSS and CROSS test). Haplotype analysis makes use of LD between markers and hence extracts more information from the data on the evolutionary history of the gene than single locus analysis. Therefore, it is more powerful than single locus methods. The HSS and CROSS test already showed their validity both in simulation studies as well as in empirical data.[14;17-19] Recently, we published a study on the same SLE population as used here on the candidate region chromosome 1, using the same tests.[16] Furthermore, a major advantage of this study is that

its population consists of exclusively Caucasian subjects from a stable founder population.
In conclusion, in a stable founder, Caucasian population of SLE patients, a strong association with HLA class I, II and III was found. Furthermore, we were also able to demonstrate that only HLA markers of class I and II are independently associated with SLE.

References

1. Tsao BP. Update on human systemic lupus erythematosus genetics. Curr.Opin.Rheumatol. 2004;16:513-521.
2. Kelly JA, Moser KL, Harley JB. The genetics of systemic lupus erythematosus: putting the pieces together. Genes Immun. 2002;3 Suppl 1:S71-S85.
3. Croker JA, Kimberly RP. Genetics of susceptibility and severity in systemic lupus erythematosus. Curr.Opin.Rheumatol. 2005;17:529-537.
4. Graham RR, Ortmann WA, Langefeld CD et al. Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. Am.J.Hum.Genet. 2002;71:543-553.
5. Smerdel-Ramoya A, Finholt C, Lilleby V et al. Systemic lupus erythematosus and the extended major histocompatibility complex--evidence for several predisposing loci. Rheumatology.(Oxford) 2005;44:1368-1373.
6. Rood MJ, van Krugten MV, Zanelli E et al. TNF-308A and HLA-DR3 alleles contribute independently to susceptibility to systemic lupus erythematosus. Arthritis Rheum. 2000;43:129-134.
7. Tsuchiya N, Kawasaki A, Tsao BP et al. Analysis of the association of HLA-DRB1, TNFalpha promoter and TNFR2 (TNFRSF1B) polymorphisms with SLE using transmission disequilibrium test. Genes Immun. 2001;2:317-322.
8. Harley JB, arcon-Riquelme ME, Criswell LA et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat.Genet. 2008;40:204-210.
9. Fernando MM, Stevens CR, Sabeti PC et al. Identification of Two Independent Risk Factors for Lupus within the MHC in United Kingdom Families. PLoS.Genet. 2007;3:e192.
10. Wallace DJ, Hahn BH. Dubois' lupus erythematosus. Philadelphia: Lippincott Williams & Wilkins; 2002.
11. Lee YH, Nath SK. Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. Hum.Genet. 2005;118:434-443.
12. Tan EM, Cohen AS, Fries JF et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1982;25:1271-1277.
13. Beckmann L, Fischer C, Deck KG et al. Exploring haplotype sharing methods in general and isolated populations to detect gene(s) of a complex genetic trait. Genet.Epidemiol. 2001;21 Suppl 1:S554-S559.
14. Boon M, Nolte IM, Bruinenberg M et al. Mapping of a susceptibility gene for multiple sclerosis to the 51 kb interval between G511525 and D6S1666 using a new method of haplotype sharing analysis. Neurogenetics. 2001;3:221-230.
15. Levinson DF, Kirby A, Slepner S et al. Simulation studies of detection of a complex disease in a partially isolated population. Am.J.Med.Genet. 2001;105:65-70.
16. Martens HA, Nolte IM, van der SG et al. Association of poly(ADP-ribose) polymerase 1 and a novel candidate locus, LOC127086, with systemic lupus erythematosus. Ann. Rheum.Dis. 2007;66:424-425.
17. Sham PC, Curtis D. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. Ann.Hum.Genet. 1995;59:323-336.
18. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am.J.Hum.Genet. 1993;52:506-516.

19. Spijker GT, Nolte IM, Jansen RC, te Meerman GJ. Genetic association studies in complex disease: disentangling additional predisposing loci from associated neutral loci using a constrained - permutation approach. Ann.Hum.Genet. 2005;69:90-101.
20. McHugh NJ, Owen P, Cox B, Dunphy J, Welsh K. MHC class II, tumour necrosis factor alpha, and lymphotoxin alpha gene haplotype associations with serological subsets of systemic lupus erythematosus. Ann.Rheum.Dis. 2006;65:488-494.
21. Diamond B, Katz JB, Paul E et al. The role of somatic mutation in the pathogenic anti-DNA response. Annu.Rev.Immunol. 1992;10:731-757.
22. Graham RR, Ortmann W, Rodine P et al. Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. Eur.J.Hum.Genet. 2007;15:823-830.
23. Gambelunghe G, Gerli R, Bocci EB et al. Contribution of MHC class I chain-related A (MICA) gene polymorphism to genetic susceptibility for systemic lupus erythematosus. Rheumatology.(Oxford) 2005;44:287-292.
24. Morgan BP, Walport MJ. Complement deficiency and disease. Immunol.Today 1991;12:301-306.
25. Pickering MC, Walport MJ. Links between complement abnormalities and systemic lupus erythematosus. Rheumatology.(Oxford) 2000;39:133-141.
26. Studnicka-Benke A, Steiner G, Petera P, Smolen JS. Tumour necrosis factor alpha and its soluble receptors parallel clinical disease and autoimmune activity in systemic lupus erythematosus. Br.J.Rheumatol. 1996;35:1067-1074.
27. Lee YH, Harley JB, Nath SK. Meta-analysis of TNF-alpha promoter -308 A/G polymorphism and SLE susceptibility. Eur.J.Hum.Genet. 2006;14:364-371.
28. Diepstra A, Niens M, Vellenga E et al. Association with HLA class I in Epstein-Barr-virus-positive and with HLA class III in Epstein-Barr-virus-negative Hodgkin's lymphoma. Lancet 2005;365:2216-2224.